Data Analysis Fundamentals

By

Dave Wells

Mark Peco

Module 0: About the Course (10 mins)

Module 1: Introduction to Data Analysis (54 mins)
- *What is Data Analysis?*
  - Data Analysis Defined
  - Finding Patterns in Data
  - Finding Meaning in Data
  - Finding Meaning in Data – Asking Questions
  - Finding Meaning in Data – Forming Hypothesis
  - Quantitative vs. Qualitative
  - Qualitative Analysis Example
  - Analysis and Analytics
  - Goals of Data Analysis
- *Kinds of Data Analysis?*
  - Statistical Analysis
  - Statistical Analysis – Statistics to Describe and Infer
  - Business Analysis
  - Business Analysis – BI, Analytics, and Data Science
  - Business Intelligence – Query and Reporting
  - Business Intelligence – OLAP
  - Business Intelligence – Slice & Dice with OLAP
  - Business Analytics
  - Business Analytics – Four Kinds of Analytics
  - Data Science
  - Data Analysis Processes
  - Data Analysis from Needs to Value
  - Statistical Data Analysis Process
  - Performance Management Process
  - Data Science Process – CRISP-DM

Module 2. Statistics and Data Analysis (49 mins)
- *Samples and Populations*
  - Definitions
  - Examples
  - Generalizing from Data
  - The Nature & Scope of Statistics
- *Descriptive Statistics*
  - Why Descriptive Statistics?
  - Similarity
  - Variation
  - Dependencies
  - Dependencies – Visualizing Corrections
  - Shape
  - Distribution
  - Common Distributions
  - Normal Distribution
- *Inferential Statistics*
  - From Descriptive to Inferential Statistics
  - Probability
  - From Descriptive to Inferential Statistics with Probabilities

- Random Variables
- Inferential Statistics Applications
- Estimating
- Generalizing
- Hypothesis Testing
- Prediciting
- *A Statistical Problem Example*
- *Framing a Statistical Problem*
- *The Descriptive Statistics*
- *Drawing Inference*

Module 3. Project Framing and Data Acquisition (37 mins)
- *Project Framing*
  - Project Framing Questions
  - Project Framing Processes
- *Problem Framing*
  - Problem Framing – Key Questions
  - Problem Framing – Problem Kernel
  - Problem Framing – Influence Mode
  - Problem Framing – Causal Model
- *Searching for Data*
  - Searching for Data – Data Requirements
  - Searching for Data – Crowdsourcing
  - Searching for Data – Data Catalog
- *Acquiring Data*
  - Choosing Datasets
  - Acquiring Data – The Questions
  - Acquiring Data – The Methods
  - Acquiring Data – The Choices

Module 4. Data Exploration and Preparation (55 mins)
- *Data Exploration – What and Why?*
  - Data Exploration Context
  - Data Exploration – Understanding the Context
- *Exploring with Data Profiling*
- Data Profiling Overview
  - Column Profiling
  - Table Profiling
  - Cross-Table Profiling
  - Data Profile Analysis – Finding Patterns
  - Data Profile Analysis – Finding Dependencies
- *Exploring with Data Visualization*
  - Univariate Analysis – Distribution of Values for Quantitative Variables
  - Univariate Analysis – Distribution of Values for Categorical Variables
  - Univariate Analysis – Central Tendency & Variation
  - Bivariate Analysis – Quantitative & Quantitative
  - Bivariate Analysis – Categorical & Categorical
  - Bivariate Analysis – Categorical & Quantitative
  - Behavior over Time Analysis – Single Variable
  - Behavior over Time Analysis – Multiple Variables

- *Data Cleansing & Structuring*
    - Data Cleansing
    - Data Structuring – Files to Tables
    - Data Structuring – Keys and Relationships
    - Data Structuring – Tables to Files
    - Data Structuring – Data Blending
- *Data Transformations to Improve, Enrich, & Format*
    - Why Data Transformation?
    - Transforming to Improve Data
    - Transforming to Enrich Data
    - Transforming to Format Data

Module 5. Analyzing Data (50 mins)
- *Cycles of Data Analysis*
- *Statistical Data Analysis*
    - Data Sampling
    - Distribution
    - Central Tendency
    - Skew
    - Dispersion and Variability
    - Probability
    - Confidence Interval
- *Algorithmic Data Analysis*
    - What is an Algorithm?
    - Common Functions of Algorithms
    - Algorithmic Analysis – A Data Mining Example
    - Algorithms and Data Science – Analytic Modeling
- *Data Visualization*
    - Data Visualization Functions
    - Common Charts and Graphs
    - More Charts and Graphs
    - Visual Composition
    - Visual Composition – What Do You Want to Show?
- *Data Storytelling*
    - Data Storytelling _ What and Why?
    - Data Presentation vs. Data Stories
    - Statistics vs. Stories
    - Kinds of Stories – Explanatory vs. Exploratory
    - Story Framing – The Support Structure of Stories
    - Story Crafting – Connecting Data and Narrative

Module 6. Human Factors and Data Analysis (36 mins)
- *Data Analysis and Culture*
    - Organizational Culture
    - Organizational Culture and Data Analysis
    - Analysis Culture
    - Analysis Culture and Data Analysis
- *Data Analyst Traits and Skills*
    - Innate Curiosity and Inquisitiveness
    - Collaboration

- o Critical Thinking
- o Complementary Thinking Styles
- o Apply the Thinking Styles
- *Data Analysis and Data Literacy*
  - o Data Analysis and Data Literacy Body of Knowledge